

2

OIEP

#2

RAW SEQUENCE LISTING

DATE: 08/06/2001

PATENT APPLICATION: US/09/915,873

TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

ENTERED

4 <110> APPLICANT: Stomp, Anne-Marie
 5 Dickey, Lynn
 6 Gasdaska, John
 9 <120> TITLE OF INVENTION: Expression of Biologically Active
 10 Polypeptides in Duckweed
 12 <130> FILE REFERENCE: 40989/237225
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/915,873
 C--> 14 <141> CURRENT FILING DATE: 2001-07-26
 14 <150> PRIOR APPLICATION NUMBER: US 60/293,330
 15 <151> PRIOR FILING DATE: 2001-05-23
 17 <150> PRIOR APPLICATION NUMBER: US 60/221,705
 18 <151> PRIOR FILING DATE: 2000-07-31
 20 <160> NUMBER OF SEQ ID NOS: 8
 22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 554
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Zea mays
 29 <400> SEQUENCE: 1
 30 gatcaagtgc aaaggtccgc cttgtttctc ctctgtctct tgatctgact aatcttggtt 60
 31 tatgattcgt tgagtaattt tggggaaagc ttcgtccaca gttttttttt cgatgaacag 120
 32 tgccgcagtg gcgctgatct tgtatgctat cctgcaatcg tggatgaactt atgtctttta 180
 33 tatccttcac taccatgaaa agactagtaa tctttctcga tgtaacatcg tccagcactg 240
 34 ctattaccgt gtggtccatc cgacagctctg gctgaacaca tcatacgata ttgagcaaag 300
 35 atctatcttc cctgttcttt aatgaaagac gtcattttca tcagtatgat ctaagaatgt 360
 36 tgcaacttgc aaggaggcgt ttctttcttt gaatttaact aactcgttga gtggccctgt 420
 37 ttctcggacg taaggccttt gctgctccac acatgtccat tcgaatttta ccgtgttttag 480
 38 caagggcgaa aagtttgcat cttgatgatt tagcttgact atgcgattgc tttcctggac 540
 39 ccgtgcagct gcgg 554
 41 <210> SEQ ID NO: 2
 42 <211> LENGTH: 498
 43 <212> TYPE: DNA
 44 <213> ORGANISM: Artificial Sequence
 46 <220> FEATURE:
 47 <223> OTHER INFORMATION: Duckweed codon optimized nucleotide sequence
 48 encoding human alpha-2B interferon
 50 <221> NAME/KEY: CDS
 51 <222> LOCATION: (1)...(498)
 53 <400> SEQUENCE: 2
 54 tgc gac ctc ccc cag acc cac agc ctc ggg tcc cgc cgc acc ctc atg 48
 55 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met
 56 1 5 10 15
 58 ctg ctg gcg cag atg cgc cgc atc tcg ctc ttc agc tgc ctg aag gac 96
 59 Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp
 60 20 25 30
 62 cgc cac gac ttc ggc ttc ccg cag gag gag ttc ggc aac cag ttc cag 144
 63 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,873

DATE: 08/06/2001

TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

```

64          35          40          45
66 aag gcc gag acg atc ccc gtg ctc cac gag atg atc cag cag atc ttc 192
67 Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile Phe
68          50          55          60
70 aac ctg ttc agc acc aag gac agc tcg gcc gcc tgg gac gag acc ctg 240
71 Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr Leu
72 65          70          75          80
74 ctc gac aag ttc tac acc gag ctg tac cag cag ctc aac gac ctg gag 288
75 Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu
76          85          90          95
78 gcg tgc gtg atc cag ggg gtt ggg gtt acg gag acg ccg ctg atg aag 336
79 Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met Lys
80          100          105          110
82 gag gac agc atc ctc gcc gtg cgc aag tac ttc cag cgc atc acg ctc 384
83 Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr Leu
84          115          120          125
86 tac ctc aag gag aag aag tac agc ccg tgc gcc tgg gag gtc gtt cgc 432
87 Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg
88          130          135          140
90 gcc gag atc atg cgc tcc ttc agc ctg agc acc aac ctc cag gag agc 480
91 Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu Ser
92 145          150          155          160
94 ctc cgc tcc aag gag taa 498
95 Leu Arg Ser Lys Glu *
96          165
99 <210> SEQ ID NO: 3
100 <211> LENGTH: 96
101 <212> TYPE: DNA
102 <213> ORGANISM: Oryza sativa
104 <400> SEQUENCE: 3
105 accatgcagg tcttgaacac gatggtcaac aagcacttcc tctccctgtc cgctctcatc 60
106 gtctctctcg ggctgagcag caacctcacc gccggc 96
108 <210> SEQ ID NO: 4
109 <211> LENGTH: 188
110 <212> TYPE: PRT
111 <213> ORGANISM: Homo sapiens
113 <400> SEQUENCE: 4
114 Met Ala Leu Thr Phe Ala Leu Leu Val Ala Leu Leu Val Leu Ser Cys
115 1          5          10          15
116 Lys Ser Ser Cys Ser Val Gly Cys Asp Leu Pro Gln Thr His Ser Leu
117          20          25          30
118 Gly Ser Arg Arg Thr Leu Met Leu Leu Ala Gln Met Arg Arg Ile Ser
119          35          40          45
120 Leu Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
121          50          55          60
122 Glu Phe Gly Asn Gln Phe Gln Lys Ala Glu Thr Ile Pro Val Leu His
123 65          70          75          80
124 Glu Met Ile Gln Gln Ile Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser
125          85          90          95

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/915,873

DATE: 08/06/2001

TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

```

126 Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr
127          100          105          110
128 Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Gly Val Gly Val
129          115          120          125
130 Thr Glu Thr Pro Leu Met Lys Glu Asp Ser Ile Leu Ala Val Arg Lys
131          130          135          140
132 Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Lys Glu Lys Lys Tyr Ser Pro
133 145          150          155          160
134 Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu
135          165          170          175
136 Ser Thr Asn Leu Gln Glu Ser Leu Arg Ser Lys Glu
137          180          185
140 <210> SEQ ID NO: 5
141 <211> LENGTH: 165
142 <212> TYPE: PRT
143 <213> ORGANISM: Homo sapiens
145 <400> SEQUENCE: 5
146 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met
147 1          5          10          15
148 Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp
149          20          25          30
150 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln
151          35          40          45
152 Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile Phe
153          50          55          60
154 Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr Leu
155 65          70          75          80
156 Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu
157          85          90          95
158 Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met Lys
159          100          105          110
160 Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr Leu
161          115          120          125
162 Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg
163          130          135          140
164 Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu Ser
165 145          150          155          160
166 Leu Arg Ser Lys Glu
167          165
170 <210> SEQ ID NO: 6
171 <211> LENGTH: 31
172 <212> TYPE: PRT
173 <213> ORGANISM: Oryza sativa
175 <400> SEQUENCE: 6
176 Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser
177 1          5          10          15
178 Val Leu Ile Val Leu Leu Gly Leu Ser Ser Asn Leu Thr Ala Gly
179          20          25          30
182 <210> SEQ ID NO: 7

```

RAW SEQUENCE LISTING

DATE: 08/06/2001

PATENT APPLICATION: US/09/915,873

TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

```

183 <211> LENGTH: 31
184 <212> TYPE: PRT
185 <213> ORGANISM: Artificial Sequence
187 <220> FEATURE:
188 <223> OTHER INFORMATION: Modified rice alpha-amylase signal peptide
190 <400> SEQUENCE: 7
191 Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser
192 1          5          10          15
193 Val Leu Ile Val Leu Thr Val Leu Ser Ser Asn Leu Thr Ala Gly
194          20          25          30
197 <210> SEQ ID NO: 8
198 <211> LENGTH: 21
199 <212> TYPE: PRT
200 <213> ORGANISM: Arabidopsis thaliana
202 <400> SEQUENCE: 8
203 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
204 1          5          10          15
205 Leu Ser Ser Ala Glu
206          20

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/915,873

DATE: 08/06/2001

TIME: 15:34:24

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date